DOCUMENT RESUME

ED 161 935 TM 007 964

AUTHOR Shiffler, Nancy L.; And Others
TITLE FEHR-Practicum: A Computer-Simulation Approach to Teaching Research and Evaluation Methods.
PUB DATE Mar 78
CONTRACT OEG-0-72-0529
NOTE 24p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978); For related documents, see ED 131 055 and 151 364-366

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Academic Achievement; *Computer Assisted Instruction; *Course Evaluation; *Educational Research; Evaluation Methods; Formative Evaluation; Higher Education; *Methods Courses; *Practicums; Research Methodology; *Research Skills; Simulation; Student Attitudes
IDENTIFIERS *Formative Evaluation Heuristic Research Practicum

ABSTRACT

The Formative Evaluation and Heuristic Research (FEHR) Practicum, a computer-simulated educational research and evaluation experience, was assessed as a means of improving participants' knowledge and skill in traditional research. Trainee activities include the preparation of preliminary surveys, evaluation proposals, budgets, computer commands, and final reports. It was hypothesized that achievement in a two-semester graduate-level course in research design and data analysis would improve in direct relation to the amount of exposure to FEHR. During the first semester the effects of FEHR on final exam scores and on perceived research competence and interest in research were assessed; the Self Assessment of Research and Evaluation Skills was used to measure attitudes toward research. The second-semester evaluation examined the effects of differing amounts of exposure to FEHR on the development of applied research skills; that is, students' FEHR project reports were rated on a five-part criterion-referenced scale. Evaluation data from both semesters supported the hypothesis that increased FEHR exposure would produce a monotonic increase in achievement in basic statistics and research design and in the applied skills represented by the final report. The attitudinal measures provide less consistent support for the hypothesized monotonic trend. The FEHR-Practicum Rating Sheet (used in this study to rate proposals and final reports) is appended. (Author/CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED161935

Session II 4.04

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Nancy L. Shiffler

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

FEHR-PRACTICUM: A COMPUTER-SIMULATION APPROACH TO TEACHING RESEARCH AND EVALUATION METHODS

by

Nancy L. Shiffler
LeVerne S. Collet
and
Ellen C. Armstrong

Presented at the Annual Meetings of the American Educational Research Association, Toronto, March 27-31, 1978.

TM007 964

The research reported herein was performed pursuant to contract No. OEG-0-72-0529, U.S. Office of Education.

Printed in U.S.A.

It has long been recognized that practical experience is essential to gaining competence in research and evaluation skills (e.g., Buswell and McConnell, 1966). As teachers of research we often observe that students who are involved in actual research concurrently with or immediately following a research methods course find that the latter, particularly its statistical content, makes much more sense to them because the research itself poses questions for which the methods course can provide answers. In contrast to these experiences are those of many graduate students enrolled in required educational research methods courses without concomitant practical experience, who see little relationship among statistics, practice "experiments" in textbooks, and the work for which they are being trained. Typically graduate students in education come into a research course better equipped to deal with verbal concepts than with mathematical ones, and many students exhibit considerable anxiety about the mathematical content of the course. Involvement in a research project can provide a verbal context to which mathematical concepts can be applied and can generate a need for information which increases the perceived relevance of and interest in statistical concepts.

However, access to such projects is difficult, and even when appropriate projects are available, it is almost impossible to arrange tasks to suit the educational needs of the students and the time constraints of the course. FEHR-PRACTICUM[1] is designed to provide this experience on demand, in the classroom, without the real-life risks and costs of actual research. FEHR is a computer-simulated city school system designed to expose users to many of

[1]The acronym "FEHR" stands for Formative Evaluation and Heuristic Research.

the experiences and problems confronted by real-life researchers and evaluators. Participants are "hired" by the school board to "solve" a particular problem, usually by comparing the effectiveness of the current program and one or more proposed programs. Available programs study populations ranging from pre-school to college level and include empirical evaluations of HeadStart projects, perceptual handicap remediation, remedial mathematics, busing and extended school year plans, and a teacher evaluation procedure. Descriptive details about the community are provided to make the simulation as realistic as possible; literally thousands of subjects and hundreds of variables are available to work with.

The most comprehensive version of the simulation requires trainees to conduct a preliminary survey to determine the extent of the problem, prepare a written evaluation proposal (including a budget), negotiate funding, prepare computer commands to perform the field evaluation, analyze the data, and submit a final report which includes recommendations for action. It is possible, however, to offer a scaled-down version of the simulation by reducing the number of required tasks and/or reducing the size of the available subject and variable pools.

This particular study was designed to assess the effectiveness of FEHR as a laboratory experience accompanying a graduate level research design and data analysis course. The study consisted of two phases, both using students enrolled in the two-semester sequence of the research design and data analysis course required by most doctoral programs in education at the University of Michigan. The basic purpose of Phase I was to evaluate the effects of FEHR on achivement in the content of the research course and on two attitudinal measures -- perceived research competence and interest in research. Phase II

examined the effects of differing amounts of exposure to FEHR in an applied situation -- the final reports of FEHR projects. Phase II again included an examination of the attitudinal measures used in Phase I. The design and results for Phase I will be presented first.

PHASE I DESIGN.

Subjects. The subjects consisted of students enrolled for the first term of the research design and data analysis course. As is typical of similar courses in other colleges of education only a small minority of these students had had previous research experience or a college course in mathematics.

All students attended a joint two-hour lecture session each week but were randomly assigned (with the exception of three students with schedule conflicts who had to be placed in a particular section) to one of two separate two-hour laboratory sessions. A coin toss designated one group as a control group and the other as experimental. Although the sections were initially equal in size, late registrations and one withdrawal resulted in an N of 25 for the control group and 22 for the experimental.

In addition to these two groups, there was a group of six students who happened to be enrolled simultaneously in the research design course and in a program evaluation course requiring a comprehensive FEHR project. These students were placed in the experimental section but were considered a unique group because of their additional exposure to FEHR and because the act of enrolling in the evaluation course indicated the possibility of both greater interest and greater competence in research.

Procedure. The content for which all students were responsible was presented in a two-hour lecture session at the beginning of each week. This session was attended by all students from both sections. The content of the

course was classical research design and data analysis with little emphasis on problem definition or decision-oriented research. Topics included basic research design, descriptive statistics, and inferential statistics through one-way analysis of variance.

Up until the time of the midterm exam students in both laboratory sections were given laboratory problems requiring the practical application of the principles studied in class to new data sets in laboratory handouts. To encourage generalization to practical situations, many of the problems were presented as synopsized research projects. However, no attempt was made to provide continuity of content from project to project. The instructor and his teaching assistant circulated about the laboratory helping students complete their assigned exercises.

Following the midterm, the experimental treatments were begun in the laboratory sessions. These are described below under three headings: control, experimental restricted, and experimental intensive.

1. Control. The control group continued working on laboratory handouts just as it did before the midterm, except that students were grouped into three-person teams, as was the experimental section. These groups worked on the laboratory problems collectively, but each person was expected to complete all exercises and to be able to explain what was done.
2. Experimental restricted. After the midterm the restricted group was given synopses of the seven FEHR problems and asked to form research teams of two or three members each. Each team was to complete before the final exam one restricted FEHR project for the problem of their choice. Projects were restricted in the sense that students dealt

with only one dependent variable, two or three independent variables, and only a few of the budgeting problems. No proposal or funding negotiation was required, but students carried out an evaluation using their selected variables and submitted a final report of their findings; the final reports included the problem definition, conceptual framework, methods, and results and conclusions. For the remainder of the term all laboratory sessions were devoted to the projects. Each group worked collectively on its project, but the individual members were expected to produce their own reports, to be able to explain the rationale for each step in the experiment, and to discuss the educational meaning of the findings.

3. Experimental intensive. Students who were additionally enrolled in the program evaluation course also spent the laboratory time after the midterm working on FEHR projects, but these students completed a comprehensive FEHR project, ongoing over the entire term. The comprehensive project involved submitting a formal proposal (including a complete budget) and negotiating funding, as well as completing an evaluation and writing a formal final report; no restrictions were made on variable selection. These students were expected to spend several additional hours beyond the laboratory time each week on their projects. No formal lecture was given in the evaluation course, although consulting time was available for the projects.

Instrumentation. A) Midterm Exam. The midterm, administered prior to the introduction of the FEHR treatment, consisted of 57 multiple-choice items covering elementary descriptive statistics and simple statistical inference, including independent and dependent t-tests. The exam was intended to be

used as a covariate to control for initial differences in statistical ability. The split-half reliability for the test was .83.

B) Final Exam. The final consisted of two parts: 1) 6 short-answer items covering knowledge of basic statistics and 6 brief application problems; 2) a description of a case study followed by questions requiring critical evaluation of three alternative methods of analysis. Blind grading was used to avoid bias. Each test was graded by both the instructor and the teaching assistant (correlation between gradings = .87); the mean of the two gradings was found for each item, and these means were summed to obtain the final scores.

C) Attitudinal Measures. The Self Assessment of Research and Evaluation Skills (SARES), developed by Collet et al. (1975), contained forty items describing tasks such as: "compute and interpret a one-way analysis of variance," or "distinguish among main effects, interactions, simple main effects, and confounded effects." The subject was asked to provide two ratings for each task: 1) her/his competence to perform the task, and 2) her/his interest in that task. A five-point scale was used for each rating, with one representing complete absence of the characteristic being assessed, and five representing a superior level. The two scores used for analysis were the mean perceived competence rating and the mean interest rating over the 40 items. The questionnaire was administered at the end of the term.

Phase I Hypotheses. If the FEHR system is meeting its objectives, then increases in exposure to FEHR ought to produce a monotonic increase in both achievement measures and the attitudinal measures. That is,

$G_1 < G_2 < G_3$ on all measures, where G indicates the treatment group

Results - Phase I. The term 1 final examination scores could be analyzed straightforwardly with an analysis of covariance, using the midterm exam as the covariate. The monotonicity hypothesis was tested using planned contrasts for linear and nonlinear trends. The results of this analysis are shown in Table I. As hypothesized, the means adjusted for the covariate increased monotonically: 57.47, 68.60, and 68.51 for groups 1, 2, and 3, respectively. This observed trend was supported by the indication of a significant linear trend, $t(49) = 2.35$, $p < .05$, and by a nonsignificant nonlinear trend, $t(49) = -1.48$, $p > .05$. A post hoc Newman-Keuls analysis indicated a significant difference between the control and restricted group means but not between the restricted and the intensive group means.

A similar analysis of covariance followed by planned contrasts was performed on the two attitudinal measures -- perceived competence and interest -- again using the midterm exam as a covariate. (As it turned out, however, the midterm was not significantly related to the attitudinal measures.) Results for perceived competence (Table II) supported the hypothesized monotonic trend with a significant linear trend, $t(49) = 2.72$, $p < .01$, and a nonsignificant nonlinear trend, $t(49) = 1.89$, $p > .05$. The adjusted means were 2.77, 2.79, and 3.54 for groups 1, 2, and 3, respectively. Newman-Keuls comparisons indicated a significant difference between the restricted and intensive group means but not between the control and restricted group means. The monotonicity hypothesis was not supported for the interest variable (Table III). The linear trend was not significant, $t(49) = .65$, $p > .05$, while the nonlinear trend was significant, $t(49) = 2.35$, $p < .05$. The adjusted means in this case were 3.70, 3.30, and 3.90 for groups 1, 2, and 3. Newman-Keuls comparisons indicated a significant difference only between the restricted and the intensive group means.

PHASE II DESIGN.

Subjects. Forty-two of the 53 first term students enrolled in the second term course -- 15 (out of 25) from the control group, 21 (out of 22) from the restricted group, and 6 (out of 6) from the intensive group. As in the first term, some students (not necessarily the same ones) were simultaneously enrolled in a program evaluation course. In addition to the students who had enrolled the previous term there was also a set of 16 new students not enrolled the previous term but who had taken at some time an equivalent course; 11 of these students were enrolled only in the research design course, and 5 were additionally enrolled in the evaluation course.

Procedure. The second phase of the study was designed to compare the effects of differing levels of involvement with FEHR on the development of applied research skills as represented by the FEHR final reports. In the second term of the course sequence, both lab sections completed a comprehensive FEHR project; there was no control laboratory section. As before, there was a common lecture session once a week covering intermediate research design, factorial analysis of variance, analysis of covariance, multiple regression, non-parametric statistics, and an introduction to multivariate techniques. Laboratory time was spent working on the FEHR projects, which required the same tasks as those of the first term intensive group. Students who were also enrolled in the evaluation course were required to complete a project different from the one they were using in the research design course (i.e., each student completed two comprehensive projects).

The two types of dependent variables, achievement and attitude, each required slightly different groups in Phase II; consequently, the groupings and results will be presented separately for each.

Final Report Groupings. Phase I was concerned with a comparison of final reports for both terms; it is important to note that the groupings defined below are groups of reports rather than groups of people. Available reports represented the following types of experiences: 1) the restricted experience from term 1; 2) the intensive experience, i.e., a single comprehensive project -- these reports were the term 1 intensive group reports and the term 2 reports of all students who had not previously or concurrently completed another comprehensive project; 3) the extensive-intensive experience -- these reports were from students who had completed 2 or more comprehensive projects, i.e., the term 2 reports from those students who had taken the program evaluation course in either or both of the terms. These groupings produce repeated measures on all but the new students; the problems involved with the analysis of this unbalanced type of design are discussed in the results section.

Final Report Measures - The Product Rating Sheet. The Product Rating Sheet was originally designed to provide an objective measure of the quality of proposals and final reports from a broad range of empirical investigations. The strategy of the rating sheet is to identify elements or characteristics of proposals or reports that are unequivocally identifiable and generally desirable. Development of the organization and general content of this instrument was heavily dependent on the work of Resta and Baker (1972) and Tuckman (1972), but specific criteria and weightings were derived by Collet et al., (1975).

A copy of the rating sheet appears as Appendix A. The rater's task is to enter a 0-to-maximum rating for each item; maximum ratings reflect importance weightings of items. For this study five subscales were used based

on items relevant and common to the assigned tasks for each group: 1) Introduction and problem definition -- items in part B1; 2) Conceptual framework -- items in part B3; 3) Method -- items in part B4; 4) Results and conclusions -- items in parts B8 and B9; 5) General Evaluation -- items in part B10.

Final Report Hypotheses. As with Phase I, a monotonic increase in Final Report ratings with increased exposure to FEHR was hypothesized:

$G_1 < G_2 < G_3$ for each subscale of the Product Rating Sheet.

Final Report Results. Phase II analyses presented considerably more difficulties than Phase I because of the unbalanced repeated measures nature of the design, the addition of new students during the second term, and the multivariate procedures needed to assess the five final report ratings. Because techniques were not readily available to deal with either the unbalanced repeated measures or with a multivariate repeated measures design, the decision was made to ignore the repeated measures aspect and simply analyze the final report ratings as a multivariate analysis of variance with follow-up tests for linearity and nonlinearity. Provided there were no test-treatment interaction, this ought to be the more conservative procedure. Because of these analysis problems, our intent was to use the statistical results only as support for a less stringent "logical" analysis of the patterns of means, i.e., our basic interest was in whether or not there was a consistent monotonic pattern in the means.

The final reports of the new students in the intensive and extensive-intensive groups were initially analyzed separately from their original sample counterparts (i.e., the data were analyzed as a 5-group MANOVA rather than a 3-group MANOVA). As they did not differ significantly from the groups made up of original students, the reports of the new and old students were

combined for subsequent analysis, the results of which appear in Table IV. An overall difference among the three groups (restricted, intensive, extensive-intensive) was found ($F(10,166) = 5.725$, $p < .01$), and the Scheffé analysis of trends using a unit weight on each of the five rating sheet subscales indicated a significant ($p < .01$) linear trend and a nonsignificant ($p > .05$) curvilinear trend. Scheffé analyses of trends by variable indicated significant linear trends for the Introduction and Problem Definition section and for the General Evaluation section. No significant nonlinear trends were found. An examination of the means shows that for each of the five subscales there was a consistent monotonic increase in ratings as exposure to FEHR increased.

Phase II Attitudinal Score Groupings. Comparisons for the attitudinal measures involved groupings parallel but not identical to those for the final reports, and also allowed for the inclusion of the term 1 control group (again, groupings will be more understandable if one thinks in terms of groups of scores rather than groups of people): 1) scores from the term 1 control group; 2) scores from both term 1 experimental groups; 3) the term 2 scores for the term 2 intensive group (i.e., all intensive subjects except those 6 from the term 1 intensive group); 4) the term 2 scores for the extensive-intensive group. The inclusion of the term 1 intensive group with the term 1 restricted group rather than with the term 2 intensive group was done because perceived competency was expected to change from term 1 and term 2 as new course material was introduced.

Attitudinal Measures. The measures of perceived competency and interest used in Phase I were used again in Phase II.

Phase II Attitudinal Hypotheses. A monotonic increase in perceived competence and interest in research with increased exposure to FEHR was

hypothesized:

$G_1 < G_2 < G_3 < G_4$ for both attitudinal measures.

Phase II Attitudinal Results. The two attitudinal measures -- perceived competence and interest -- were analyzed using an analysis of variance for the four groups defined above (term 1 control, term 1 experimental, term 2 intensive, and term 2 extensive-intensive). (The term 1 midterm could not be used as a covariate because it was not available for the new students; however, its relationship to the term 1 attitudinal measures was nonsignificant, so its unavailability was probably not crucial.) The results for perceived competence appear in Table V. Significant trends for both the linear ($t(96) = 2.31$, $p < .05$) and nonlinear ($t(96) = 2.27$, $p < .05$) trends were found, the latter being due to the reversal of groups three and four (term 2 intensive and term 2 extensive-intensive). The means were 2.78, 2.96, 3.28, and 3.18 for groups 1, 2, 3, and 4, respectively. Post hoc Newman-Keuls comparisons failed to produce significant differences between groups three and four, while a significant difference was found between groups one and three.

The analysis of the interest scores (Table VI) indicated only a significant curvilinear trend, $t(96) = 2.95$, $p < .01$, due to the reversal between groups one and two. The means were 3.68, 3.44, 3.83, and 4.18 for groups 1, 2, 3, and 4. Post hoc Newman-Keuls comparisons indicated that groups one and two did not differ significantly, while group four differed from both group one and group two. Group three did not differ significantly from the other groups.

Discussion. The data from this evaluation uniformly support the hypothesis that increased FEHR exposure would produce a monotonic increase in achievement in basic statistics and research design, and in the applied skills represented

by the final reports. Although, as mentioned above, the Phase II statistical procedures were not completely adequate, the pattern of results taken as a whole appears persuasive. From a statistical and design standpoint, the "purest" comparison was that between the control and restricted groups from term 1; random assignment to treatment ensured the initial comparability of these two groups, while self-selection occurred for the term 1 intensive group, whose members elected to take the additional program evaluation course. The relatively large downward adjustment of the final test mean given the midterm covariate for the intensive group suggests that this group may have been initially better prepared for the research course. However, because for this group the FEHR treatment was ongoing over the entire term, the midterm covariate may have been removing some of the treatment effect for this group. Thus, the control vs. restricted comparison directly supported the hypothesized advantage for FEHR, while the results for the intensive group lent additional support.

Phase II involved a similar self-selection process as well as the unbalanced repeated mesaures problem, but the consistency of the results provides support for the the hypotheses regarding applied skills.

The attitudinal measures provide less consistent support for the hypothesized monotonic trend, but neither do they completely contradict it. The perceived competency scores show a generally upward trend with increased exposure to FEHR, but it could be argued that the greatest increases in this variable occur at points where they would be expected in any case -- between the restricted and the "self-selected" intensive group in Phase I and between the term 1 (i.e., groups 1 and 2) and term 2 (i.e., groups 3 and 4) scores in Phase II. The decrease between the intensive and the extensive-intensive groups in Phase II, although nonsignificant, has occurred occasionally

in other uses of FEHR, and it has been suggested that as one becomes more aware of the complexity of research, perceptions of one's relative competence to deal with it decrease. However, a more direct study of the problem would be necessary to test this idea adequately.

A curvilinear trend was found for the interest scores, with the restricted group showing somewhat less interest than the control group. However, as exposure increases beyond the restricted group, interest increases again. Our other experiences with FEHR have indicated that many students initially have somewhat unrealistic expectations about what can be done with statistics and therefore have a high interest based on these unrealistic expectations. Involvement in research produces more realistic expectations resulting at first in a decrease in interest level; as involvement increases, interest begins to increase again, but this time it is based on a more realistic outlook.

Overall, the results of this study support the notion that involvement in simulated research enhances learning in traditional research design and analysis courses and that the more "research" the student does the greater the learning. The FEHR-PRACTICUM has the capacity to provide this experience within the time constraints and the didactic needs of the course and without the risks or access difficulties of real-world research.

TABLE I. ANALYSIS OF COVARIANCE OF THE TERM 1 FINAL EXAM SCORES USING THE MIDTERM AS COVARIATE

Source	DF	SUM OF SQUARES	MEAN SQUARE	F-STAT.	SIGNIF.
Equality of Adj. Cell Means	2	1672.7	836.36	6.854	.002
Zero Slope	1	2639.5	2639.5	21.631	.000
Error	49	5979.3	122.03		
Equality of Means (w/o Covariates)	2	2330.3	1165.2	6.759	.002
Error	50	8618.8	172.38		

TABLE OF COEFFICIENTS

Covariate	COEFF.	STD. ERROR	T-STAT.	SIGNIF.
Midterm	1.103	.237	4.650	.000

TABLE OF MEANS

Exposure Level	(1) Control	(2) Restricted	(3) Intensive
Mean	56.600	68.545	73.333
Adj. Mean	57.470	68.600	69.508
(Std. Error)	2.217	2.355	4.584
Sample Size	25	22	6

COMPARISON TESTED	VALUE OBSERVED	T-STAT.	SIGNIF.
Linear Trend	12.039	2.350	.022
Curvilinear Trend	-10.223	-1.478	.145

TABLE II. PHASE I ANALYSIS OF COVARIANCE OF RESEARCH COMPETENCE USING THE MIDTERM AS COVARIATE

Source	DF	SUM OF SQUARES	MEAN SQUARE	F-STAT.	SIGNIF.
Equality of Adj. Cell Means	2	2.976	1.488	3.973	.025
Zero Slope	1	.901	.901	2.405	.127
Error	49	18.354	.374		
Equality of Means (w) Covariate)	2	2.476	1.236	3.210	.048
Error	50	19.255	.385		

TABLE OF COEFFICIENTS

Covariate	COEFF.	STD. ERROR	T-STAT.	SIGNIF.
Midterm	-.020	.013	-1.551	.127

TABLE OF MEANS

Exposure Level	(1) Control	(2) Restricted	(3) Intensive
Mean	2.783	2.793	3.469
Adj. Mean	2.767	2.792	3.540
(Std. Error)	.122	.130	.253
Sample Size	25	22	6

COMPARISON TESTED	VALUE OBSERVED	T-STAT.	SIGNIF.
Linear Trend	.772	2.723	.008
Curvilinear Trend	.723	1.887	.065

TABLE III. PHASE I ANALYSIS OF COVARIANCE OF INTEREST USING THE MIDTERM AS COVARIATE

Source	DF	SUM OF SQUARES	MEAN SQUARE	F-STAT.	SIGNIF.
Equality of Adj. Cell Means	2	2.636	1.318	2.864	.066
Zero Slope	1	.763	.763	1.659	.203
Error	49	22.545	.460		
Equality of Means (w/o Covariate)	2	2.852	1.426	3.059	.055
Error	50	23.308	.466		

TABLE OF COEFFICIENTS

Covariate	COEFF.	STD. ERROR	T-STAT.	SIGNIF.
Midterm	.018	.014	1.288	.203

TABLE OF MEANS

Exposure Level	(1) Control	(2) Restricted	(3) Intensive
Mean	3.684	3.301	3.968
Adj. Mean	3.699	3.302	3.903
(Std. Error)	.136	.144	.281
Sample Size	25	22	6

COMPARISON TESTED	VALUE OBSERVED	T-STAT.	SIGNIF.
Linear Trend	.204	.649	.518
Curvilinear Trend	.996	2.347	.023

TABLE IV. PHASE II RESULTS OF THE MULTIVARIATE ANALYSIS OF VARIANCE OF THE FINAL REPORT PRODUCT RATING SHEET

Equality of Group Means: DF = 10,166.0 F = 5.725 SIGNIF. = .000

Alt. Test of Equality of Group Means: Max Root = .528 SIGNIF. = .000

TABLE OF MEANS

Exposure Level	(1) Restricted	(2) Intensive	(3) Extensive-Intensive
IP (Intro. and Problem Def.)	9.909	10.857	13.231
CF (Conceptual Framework)	8.590	9.047	11.077
M (Method)	24.364	27.929	28.308
RC (Results and Conclusions)	23.636	26.952	29.346
GE (General Evaluation)	8.454	11.857	14.192
Sample Size	22	42	26

Analysis of Trends for the Final Report Combination: Unit Weight on Each Variable

COMPARISON TESTED	OBSERVED VALUE	SCHEFFÉ CRITICAL VALUES SIG = .05	SIG = .01
Linear Trend	21.199	17.589	20.275
Curvilinear Trend	- 2.177	25.699	29.624

Linear Trend Comparison by Variable			
IP	3.321	2.559	2.950
CF	2.486	5.101	5.880
M	3.944	5.018	5.785
RC	5.709	7.948	9.762
GE	5.737	3.570	4.115
Curvilinear Trend Comparison by Variable			
IP	1.425	3.739	4.310
CF	1.572	7.454	8.592
M	-3.185	7.333	8.453
RC	- .922	11.614	13.387
GE	-1.067	5.217	6.013

TABLE V. PHASE II UNIVARIATE ANALYSIS OF VARIANCE OF PERCEIVED COMPETENCE

Source	DF	SUM OF SQUARES	MEAN SQUARE	F-STAT.	SIGNIF.
Between	3	3.787	1.262	3.006	.034
Within	96	40.315	.419		
Total	99	44.103			

Equality of Variances: DF = 3, 14778. F = .175 SIGNIF. = .912

EXPOSURE LEVEL	N	MEAN	VARIANCE	STD. DEV.
1. Control	25	2.783	.417	.646
2. Term 1 Exper.	27	2.962	.420	.686
3. Term 2 Intens.	31	3.277	.465	.682
4. Ext.-Intens.	17	3.157	.335	.579

COMPARISON TESTED	VALUE OBSERVED	T-STAT.	SIGNIF.
Linear Trend	1.047	2.313	.022
Curvilinear Trend	1.437	2.265	.025

TABLE VI. PHASE II UNIVARIATE ANALYSIS OF VARIANCE OF INTEREST

Source	DF	SUM OF SQUARES	MEAN SQUARE	F-STAT.	SIGNIF.
Between	3	6.044	2.014	4.781	.003
Within	96	40.455	.421		
Total	99	46.501			

Equality of Variances: DF = 3,14778. F = 1.565 SIGNIF. = .195

EXPOSURE LEVEL	N	MEAN	VARIANCE	STD. DEV.
1. Control	25	3.684	.505	.711
2. Term 1 Experim.	27	3.439	.509	.713
3. Term 2 Intensive	31	3.826	.399	.632
4. Ext.-Intensive	17	4.180	.193	.439

COMPARISON TESTED	VALUE OBSERVED	T-STAT.	SIGNIF.
Linear Trend	.393	.867	.387
Curvinilinear Trend	1.876	2.951	.004

REFERENCES

Buswell, G.T. and T. R. McConnell. Training for Educational Research. Center for the Study of Higher Education, Berkeley, Calif., 1966.

Collet, L.S., R. Parnes, E. Anderson, D.L. Feagans, R. Fine, M.J. Harlow, Sr. E. Rice, P. Roeper, N.L. Shiffler, B. Simon, D. Vernon. "Formative and Summative Evaluation of the FEHR-PRACTICUM Training Module." Final Report, U.S. Office of Education, Project No. 2-0035, Grant No. OEG-0-72-0529, 1975.

Resta, P.E. and R.L. Baker. Components of the Educational Research Proposal. New York: American Book Company, Van Nostrand Reinhold Company, 1972.

Tuckman, B.W. Conducting Educational Research. New York: Harcourt Brace Jovanovich, Inc., 1972.

APPENDIX A

FEHR-PRACTICUM RATING SHEET

for Proposals and Final Reports

Team_____ Members_____

Date_____ Product_____

A. Preliminary Materials

1. Title page 3 ☐
   - (a) precise prob. ident. (2)__
   - (b) suff. concise for index (1)__
   - (c) too long or wordy (-1)__
   - (d) incomplete author/info. (-1)__

2. Tables: Contents, figures, etc. 2 ☐

3. Abstract 10 ☐
   - (a) study purpose outlined (2)__
   - (b) target population identified (1)__
   - (c) major dependent variables (1)__
   - (d) design outlined (2)__
   - (e) analytic procedures outlined (2)__
   - (f) key comparisons outlined (2)__

B. Body of the Proposal

1. Introduction 20 ☐
   - (a) statement of the problem (4)__
   - (b) context or background (4)__
   - (c) purpose of proposed study (4)__
   - (d) importance of study (2)__
   - (e) scope and delimitation (2)__
   - (f) assumptions, limitations (4)__
   - (g) lacks logical relations (-8)__

2. Review of Related Literature 20 ☐
   - (a) relation articles & study (4)__
   - (b) article methods evaluated (4)__
   - (c) articles representative (4)__
   - (d) logical grouping of studies (4)__
   - (e) results summar. & synth. (4)__
   - (f) critical studies are missing (-8)__

3. Conceptual framework (rationale) 24 ☐
   - (a) set of principles or laws (4)__
   - (b) prin. tied to theory, research (4)__
   - (c) prin. form coherent unit (4)__
   - (d) prin. & modifying criteria (2)__
   - (e) research hypotheses stated (4)__
   - (f) definition of terms (2)__
   - (g) success crit. (objectives) (4)__
   - (h) lacks logical relations (-10)__

4. Method 40 ☐
   - (a) subjects are described (2)__
     - -sampling described (2)__
     - -sampling representative (4)__
   - (b) design-described (4)__
     - -rationale (2)__
     - -variables not operat. (-4)__
     - -crit. compar. groups (2)__
     - -valid comparisons (2)__
     - -inval. not controlled (-4)__
   - (inst.-desc. all tests (2)__
     - -assess rel. & val. (4)__
     - -unsuitable, incomplete (-4)__
   - (d) data-source, who admin., how (2)__
     - -data matrix defined (2)__
   - (e) analysis-rationale given (4)__
     - -covers hypotheses (4)__
     - -efficient (4)__
     - -inappropriate for purpose (-10)__

* 5. Budget 10 ☐
   - (a) source of each item clear (2)__
   - (b) standard items present (2)__
   - (c) problems anticipated (2)__
   - (d) expense resource balance (2)__
   - (e) cost effectiveness assessed (2)__

* 6. Logistics 10 ☐
   - (a) schedule of activities (2)__
   - (b) work distributed prorate (2)__
   - (c) sufficient personnel (2)__
   - (d) bottlenecks anticipated (2)__
   - (e) sequence logical & efficient (2)__

* 7. Personnel 10 ☐
   - (a) major personnel named (3)__
   - (b) personnel respons. defined (4)__
   - (c) evidence of competency (vita.) (3)__
   - (d) personnel inadequate (-10)__

**8. Results (Statistical Concl.) 30 ☐
   - (a) result for each hyp. (4)__
   - (b) explicit stat. concl. (2)__
   - (c) neat concise displays (2)__
   - (d) logical organization (6)__
   - (e) explan. graphs, diag. (4)__
   - (f) overall summary, synthesis (12)__
   - (g) procedural errors (-10)__

**9. Educational Conc. & Implic. 24 ☐
   - (a) educ. meaning results given (4)__
   - (b) obj. not subj. presentation (4)__
   - (c) pattern of results interp. (4)__
   - (d) cost effectiveness assessed (8)__
   - (e) validity of concl. (target) (4)__
   - (f) misinterpretations (-10)__

10. Gen. Eval. of report/proposal 20 ☐
   - (a) physically neat and orderly (2)__
   - (b) style acceptable (AERA, etc.) (5)__
   - (c) appropriate citations given (3)__
   - (d) organization clear, readable (5)__
   - (e) study is replicable (5)__

C. Supplementary Materials (bonus) 15 ☐
   1. Bibliography (5)__
   2. Appended explanations of data (10)__

MAX. POSSIBLE/TOTAL _____/☐

RATING_____

* Applicable to Proposal only

** Applicable to Final Report only